

Schémas de lifting adaptatifs via des critères parcimonieux

Mounir KAANICHE¹, Jean-Christophe PESQUET², Amel BENAZZA-BENYAHIA³, Béatrice PESQUET-POPESCU¹

¹Département Traitement du Signal et des Images, Télécom ParisTech
46 rue Barrault, 75634 Paris Cédex 13, France

²Laboratoire d'Informatique Gaspard Monge, Université Paris-Est
Champs-sur-Marne, 77454 Marne-la-Vallée, France

³Unité de Recherche en Imagerie Satellitaire et ses Applications, SUP'COM, Univ. de Carthage
Cité Technologique des Communications, 2083, Tunisie

kaaniche@telecom-paristech.fr, jean-christophe.pesquet@univ-paris-est.fr
benazza.amel@supcom.rnu.tn, beatrice.pesquet@telecom-paristech.fr

Résumé – Dans cet article, nous nous intéressons à l'optimisation des filtres de prédiction et de mise à jour apparaissant dans un schéma de lifting 2D non-séparable par des méthodes d'optimisation convexe non lisse. Comparativement à l'approche usuelle reposant sur la minimisation de la norme euclidienne ℓ_2 , la méthode proposée permet de minimiser des critères parcimonieux. Plus précisément, nous proposons de nouveaux algorithmes efficaces pour minimiser des critères de type ℓ_1 et ℓ_1 pondéré.

Abstract – In this paper, we investigate different optimization techniques for the design of all the filters defined in a 2D non-separable lifting structure. Indeed, unlike the most commonly used approach based on the minimization of an ℓ_2 - norm, the proposed method aims at minimizing new sparsity criteria. More precisely, we focus on the use of ℓ_1 and weighted ℓ_1 criteria.

1 Introduction

Depuis les travaux de Sweldens [1], les schémas de lifting, appelés également transformations en ondelettes de seconde génération, ont connu un grand succès. Ces structures sont des outils bien appropriés à la représentation d'une image sur plusieurs niveaux de résolution, tout en assurant de façon intrinsèque la propriété de reconstruction parfaite. De telles caractéristiques sont appréciées en compression d'images dans des applications de transmission progressive d'images sur des réseaux à qualité variable. Dans son principe de base, un schéma de lifting appliqué à un signal 1D, peut être décomposé en trois parties. Tout d'abord, une transformée polyphase est effectuée permettant de séparer le signal en deux sous-ensembles disjoints formés respectivement des échantillons d'indices pairs et d'indices impairs. Ensuite, une prédiction suivie d'une opération de mise à jour sont appliquées afin de générer un signal de détails et un signal d'approximation. Généralement, ces schémas 1D sont étendus au cas 2D de manière séparable : à chaque niveau de résolution, le lifting s'applique aux lignes puis aux colonnes (ou *vice-versa*) conduisant ainsi à une sous-bande d'approximation et à trois sous-bandes de coefficients de détails orientés horizontalement, verticalement et diagonalement. Les performances des schémas de codage peuvent ainsi dépendre fortement des opérateurs de lifting (filtres de prédiction et de mise à jour) mis en jeu. C'est pourquoi une grande attention a été portée à l'optimisation de ces opérateurs afin de concevoir des schémas de lifting mieux *adaptés* au *contenu* du

signal d'entrée et fournissant des représentations compactes. La plupart des méthodes développées dans la littérature se sont intéressées à l'optimisation des filtres de prédiction [2, 3, 4]. Notamment, certains travaux ont proposé de construire le filtre de prédiction en minimisant l'entropie des signaux de détails [4]. Cependant, l'entropie du signal de détail étant une fonction non linéaire implicite des coefficients du filtre de prédiction à optimiser, sa minimisation ne peut se faire que par des algorithmes itératifs. En particulier, dans [4], l'algorithme de Nelder-Mead a été utilisé pour minimiser l'entropie des signaux de détails. Toutefois, ces algorithmes d'optimisation présentent quelques inconvénients. Tout d'abord, ils dépendent des conditions initiales et peuvent converger vers des minima locaux. De plus, ils sont très lourds en temps de calcul. Pour remédier à ces problèmes, une alternative reposant sur la minimisation de la variance du signal de détail a également été envisagée [2, 3]. En revanche, le problème d'optimisation du filtre de mise à jour a été relativement peu abordé. La méthode conventionnelle consiste à optimiser le filtre de mise à jour en minimisant l'erreur de reconstruction [2]. Notons que la flexibilité dans le choix des opérateurs du lifting a permis de développer des schémas de lifting directionnels adaptatifs [3, 5, 6]. Cet article est structuré comme suit. La section 2 décrit la structure de lifting considérée ainsi que les approches d'optimisation de ses différents opérateurs. Nous détaillons ensuite dans la section 3 une nouvelle approche d'optimisation reposant sur la définition de critères parcimonieux. Des résultats expérimentaux

taux sont présentés dans la section 4. Enfin, la section 5 fournit une conclusion.

2 Structure de lifting adaptatif

2.1 Principe de la structure retenue

Soulignons qu'une structure de lifting séparable rend la décomposition peu efficace dans le cas des images présentant des contours qui ne sont ni horizontaux ni verticaux. De plus, cette contrainte de séparabilité limite le nombre de degrés de liberté dans le choix des opérateurs de prédiction et de mise à jour. C'est pourquoi nous retenons dans cet article une structure de lifting 2D non séparable composée de trois étapes de prédiction et d'une étape de mise à jour comme le montre la Fig. 1. Le principe général de cette transformation 2D non séparable

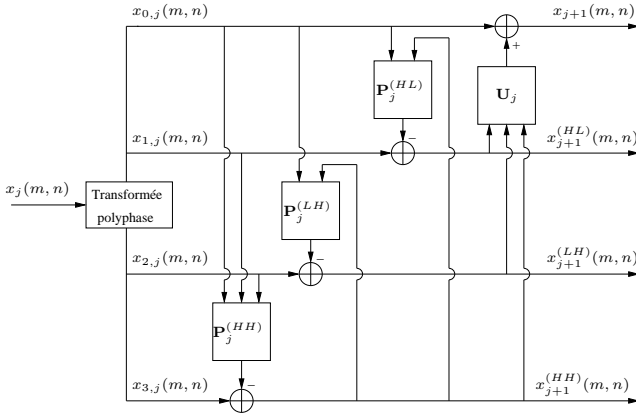


FIG. 1 – Structure de décomposition 2D non-séparable.

consiste à décomposer une image $x_j(m, n)$ en quatre composantes polyphase $x_{0,j}(m, n) = x_j(2m, 2n)$, $x_{1,j}(m, n) = x_j(2m, 2n+1)$, $x_{2,j}(m, n) = x_j(2m+1, 2n)$, et $x_{3,j}(m, n) = x_j(2m+1, 2n+1)$ pour générer une sous-bande d'approximation x_{j+1} et les sous-bandes des détails orientés horizontalement $x_{j+1}^{(HL)}$, verticalement $x_{j+1}^{(LH)}$ et diagonalement $x_{j+1}^{(HH)}$. Notons que cette structure 2D inclut comme cas particulier toute structure de lifting séparable classique où les étapes de traitement 1D sont composées d'une étape de prédiction suivie d'une étape de mise à jour (structures dites P-U).

2.2 Optimisation des filtres de prédiction et de mise à jour

Récemment, nous nous sommes intéressés à l'optimisation des filtres de prédiction et de mise à jour présents dans ces structures de lifting 2D [7]. Plus précisément, chaque filtre de prédiction $\mathbf{p}_j^{(o)}$, avec $o \in \{HL, LH, HH\}$, a été optimisé en minimisant la variance des coefficients de détails $x_{j+1}^{(o)}$. Notons que ce critère a été utilisé assez fréquemment dans la littérature en raison de sa simplicité et de son efficacité.

Concernant le choix du filtre de mise à jour, nous avons proposé une nouvelle méthode d'optimisation visant à réduire les phénomènes indésirables de repliement de spectre. Ainsi, le

filtre de mise à jour a été conçu de façon à ce que sa sortie soit la plus proche possible de celle d'un filtre passe-bas idéal. Ceci équivaut à minimiser le critère suivant :

$$\begin{aligned} \mathcal{J}(\mathbf{u}_j) &= \mathbb{E} \left[\left(x_{j+1}(m, n) - y_{j+1}(m, n) \right)^2 \right] \\ &= \mathbb{E} \left[\left(x_{0,j}(m, n) + \mathbf{u}_j^\top \mathbf{x}_{j+1}(m, n) - y_{j+1}(m, n) \right)^2 \right] \end{aligned} \quad (1)$$

où

- $\mathbf{u}_j = \left(u_j^{(o)}(k, l) \right)_{(k,l) \in \mathcal{U}_j^{(o)}, o \in \{HL, LH, HH\}}^\top$ est le vecteur des poids du filtre de mise à jour,

- $\mathbf{x}_{j+1}(m, n) = \left(x_{j+1}^{(o)}(m-k, n-l) \right)_{(k,l) \in \mathcal{P}_{i,j}^{(o)}, o \in \{HL, LH, HH\}}^\top$ est le vecteur de référence contenant les coefficients détails utilisés dans l'étape de mise à jour.

- $y_{j+1}(m, n) = (h * x_j)(2m, 2n)$ où h désigne la réponse impulsionnelle du filtre passe-bas idéal.

Bien que cette technique d'optimisation du filtre de mise à jour donne des performances en terme de qualité de reconstruction assez similaires à celles issues de la méthode conventionnelle reposant sur la minimisation de l'erreur de reconstruction, le critère proposé présente l'avantage d'être plus simple à minimiser, conduisant à un gain d'environ un facteur 3 en temps de calcul [7].

3 Critères d'optimisation proposés

Dans cet article, nous nous proposons de développer de nouveaux critères pour l'optimisation des filtres de prédiction en remplaçant le critère usuel ℓ_2 par des critères de type ℓ_1 . Une des motivations principales d'un tel choix est qu'il permet de produire des décompositions parcimonieuses (*sparse* en anglais) ayant un grand nombre de coefficients d'ondelettes nuls. Pour cela, au lieu de minimiser la variance des signaux de détails (associée à la norme ℓ_2), chaque filtre de prédiction $\mathbf{p}_j^{(o)}$, avec $o \in \{HL, LH, HH\}$, est calculé en minimisant la norme ℓ_1 des coefficients détails $x_{j+1}^{(o)}$:

$$\forall o \in \{HL, LH, HH\}, \forall i \in \{1, 2, 3\},$$

$$\begin{aligned} \mathcal{J}_{\ell_1}(\mathbf{p}_j^{(o)}) &= \sum_{m=1}^{M_{j+1}} \sum_{n=1}^{N_{j+1}} \left| x_{j+1}^{(o)}(m, n) \right| \\ &= \sum_{m=1}^{M_{j+1}} \sum_{n=1}^{N_{j+1}} \left| x_{i,j}(m, n) - (\mathbf{p}_j^{(o)})^\top \tilde{\mathbf{x}}_j^{(o)}(m, n) \right| \end{aligned} \quad (2)$$

où $x_{i,j}(m, n)$ est la $(i+1)$ ème composante polyphase à prédire, $\tilde{\mathbf{x}}_j^{(o)}(m, n)$ est un vecteur de référence contenant les échantillons utilisés dans l'étape de prédiction de $x_{i,j}(m, n)$, $\mathbf{p}_j^{(o)}$ est le vecteur de prédiction à optimiser, M_{j+1} et N_{j+1} correspondent aux dimensions de la sous-bande x_{j+1} . Notons que le critère \mathcal{J}_{ℓ_1} est non-différentiable et son optimisation peut se faire au moyen de la classe des algorithmes itératifs proximaux qui a été conçue pour résoudre des problèmes de minimisation

convexes et non nécessairement différentiables [8]. Dans ce travail, nous avons retenu l'algorithme de Douglas-Rachford dont l'efficacité dans la résolution de problèmes inverses de grande taille n'est plus à démontrer [8]. Bien que cet algorithme soit itératif, nous avons observé expérimentalement que sa convergence est généralement assurée en un faible nombre d'itérations (de l'ordre de 50 itérations). Par conséquent, la minimisation d'un critère ℓ_1 s'avère relativement peu coûteuse en temps d'exécution par rapport à l'approche usuelle reposant sur la minimisation d'un critère ℓ_2 .

Il est important de remarquer sur la Fig. 1 que le signal de détails diagonaux $x_{j+1}^{(HH)}$ est utilisé dans les deuxième et troisième étapes de prédiction pour calculer les signaux de détails horizontaux $x_{j+1}^{(HL)}$ et verticaux $x_{j+1}^{(LH)}$. Par conséquent, l'optimisation du filtre $\mathbf{p}_j^{(HH)}$ reposant uniquement sur la minimisation de la norme ℓ_1 de $x_{j+1}^{(HH)}$ est sous-optimale. Pour cette raison, il se révèle intéressant de calculer le filtre $\mathbf{p}_j^{(HH)}$ de façon à minimiser l'erreur de prédiction globale. Celle-ci s'exprime sous la forme d'une somme *pondérée* des normes ℓ_1 des sous-bandes de détails :

$$\tilde{\mathcal{J}}_{w\ell_1}(\mathbf{p}_j^{(HH)}) = \sum_{o \in \{HL, LH, HH\}} \sum_{m=1}^{M_{j+1}} \sum_{n=1}^{N_{j+1}} \sqrt{w_j^{(o)}} |x_{j+1}^{(o)}(m, n)|$$

où les $w_j^{(o)}$ sont les poids des sous-bandes des coefficients de détails $x_{j+1}^{(o)}$ obtenus en utilisant la procédure détaillée dans [9] (pour $o \in \{HL, LH, HH\}$). Pour minimiser $\tilde{\mathcal{J}}_{w\ell_1}$, nous avons utilisé l'algorithme de Douglas-Rachford formulé dans un espace produit [8]. Comme l'optimisation du filtre $\mathbf{p}_j^{(HH)}$ est effectuée en minimisant la somme pondérée des normes ℓ_1 des 3 sous-bandes détails, on peut constater que la solution optimale de $\mathbf{p}_j^{(HH)}$ va dépendre des filtres $\mathbf{p}_j^{(HL)}$ et $\mathbf{p}_j^{(LH)}$. De plus, il est clair que l'optimisation des filtres $\mathbf{p}_j^{(HL)}$ et $\mathbf{p}_j^{(LH)}$ dépend de $\mathbf{p}_j^{(HH)}$ puisque le signal $x_{j+1}^{(HH)}$ résultant de l'optimisation de $\mathbf{p}_j^{(HH)}$ est utilisé comme signal de référence dans les étapes d'optimisation de $\mathbf{p}_j^{(HL)}$ et $\mathbf{p}_j^{(LH)}$. Ainsi, on conclut que l'optimisation de $\mathbf{p}_j^{(HH)}$ dépend de $\mathbf{p}_j^{(HL)}$ et $\mathbf{p}_j^{(LH)}$, et *inversement*. Pour cela, nous proposons un algorithme itératif qui consiste à alterner entre l'optimisation des différents filtres et le calcul des poids. L'algorithme itératif est décrit de la manière suivante :

- ① Initialiser le numéro de l'itération it à 0.
 - Optimiser séparément les 3 filtres de prédiction $\mathbf{p}_j^{(HH)}$, $\mathbf{p}_j^{(LH)}$ et $\mathbf{p}_j^{(HL)}$ en minimisant respectivement $\mathcal{J}_{\ell_1}(\mathbf{p}_j^{(HH)})$, $\mathcal{J}_{\ell_1}(\mathbf{p}_j^{(LH)})$ et $\mathcal{J}_{\ell_1}(\mathbf{p}_j^{(HL)})$. Les solutions obtenues sont notées $\mathbf{p}_j^{(HH,0)}$, $\mathbf{p}_j^{(LH,0)}$ et $\mathbf{p}_j^{(HL,0)}$.
 - Optimiser le filtre de mise à jour en minimisant $\mathcal{J}(\mathbf{u}_j)$.
 - Calculer les poids $w_j^{(o,0)}$ de chaque sous-bande de détails comme décrit dans [9] et en déduire la somme pondérée des normes ℓ_1 des sous-bandes détails.
- ② pour $it = 1, 2, 3, \dots$
 - Fixer $\mathbf{p}_j^{(LH)} = \mathbf{p}_j^{(LH,it-1)}$, $\mathbf{p}_j^{(HL)} = \mathbf{p}_j^{(HL,it-1)}$, et

- optimiser $\mathbf{p}_j^{(HH)}$ en minimisant $\tilde{\mathcal{J}}_{w\ell_1}(\mathbf{p}_j^{(HH)})$. Soit $\mathbf{p}_j^{(HH,it)}$ le nouveau premier filtre de prédiction optimal.
- Optimiser $\mathbf{p}_j^{(LH)}$ en minimisant $\mathcal{J}_{\ell_1}(\mathbf{p}_j^{(LH)})$. Soit $\mathbf{p}_j^{(LH,it)}$ le nouveau deuxième filtre de prédiction optimal.
- Optimiser $\mathbf{p}_j^{(HL)}$ en minimisant $\mathcal{J}_{\ell_1}(\mathbf{p}_j^{(HL)})$. Soit $\mathbf{p}_j^{(HL,it)}$ le nouveau troisième filtre de prédiction optimal.
- Optimiser le filtre de mise à jour en minimisant $\mathcal{J}(\mathbf{u}_j)$.
- Calculer les poids $w_j^{(o,it)}$ de chaque sous-bande de détails ainsi que l'erreur de prédiction globale pondérée.

- ③ Sélectionner les filtres optimaux à partir de l'itération donnant l'erreur de prédiction globale pondérée la plus faible.

Soulignons que l'étape 3 de l'algorithme a été introduite car l'erreur de prédiction globale pondérée ne diminue pas nécessairement d'une itération à l'autre. Ceci s'explique par le fait que les poids $w_j^{(o)}$ sont recalculés à chaque itération après l'optimisation de tous les filtres.

4 Résultats expérimentaux

Pour montrer l'intérêt des critères parcimonieux proposés, nous considérons le schéma de lifting (2,2) (appelé aussi 5/3) à coefficients fixes. Plus précisément, l'équivalent de cette structure en mode non-séparable (comme l'illustre la Fig. 1) sera noté «NSLS(2,2)». Puis, les filtres de prédiction et de mise à jour sont optimisés de manière conventionnelle comme expliqué dans la section 2. Ce schéma adaptatif sera désigné par NSLS(2,2)-OPT-L2. Ensuite, en gardant la même stratégie d'optimisation du filtre de mise à jour, les filtres de prédiction sont optimisés en remplaçant la norme ℓ_2 par la norme ℓ_1 (critère \mathcal{J}_{ℓ_1}). Enfin, nous optimisons conjointement les filtres de prédiction en minimisant de façon itérative la somme pondérée des normes ℓ_1 des sous-bandes détails (critère $\tilde{\mathcal{J}}_{w\ell_1}$). Ces deux dernières méthodes seront désignées respectivement par NSLS(2,2)-OPT-L1 et NSLS(2,2)-OPT-WL1. La Fig. 3 illustre les performances de ces différentes méthodes appliquées à l'image "einst" en utilisant le codeur entropique de JPEG2000. Les résultats obtenus montrent que l'adaptation des opérateurs du lifting selon la méthode NSLS(2,2)-OPT-L2 conduit à un gain de 0.1 à 0.5 dB par rapport à NSLS(2,2). La minimisation d'un critère ℓ_1 NSLS(2,2)-OPT-L1 au lieu d'un critère ℓ_2 améliore encore les performances de 0.1 à 0.2 dB. Enfin, un gain global de 0.2 à 0.5 dB est réalisé en appliquant l'algorithme itératif permettant l'optimisation conjointe des différents filtres. Pour confirmer ces résultats, nous présentons sur la Fig. 2 les images de "lena" reconstruites à 0.125 bpp. Les performances sont évaluées en terme de PSNR et des critères perceptuels SSIM et VSNR [10]. Ainsi, il apparaît que notre approche d'optimisation améliore la qualité visuelle de l'image décodée en réalisant un gain de 0.3 dB (resp. 0.6 dB) en termes de PSNR (resp. VSNR) comparée à l'approche classique reposant sur la minimisation d'un critère quadratique.



(a) Image originale



(b) PSNR=29.86 dB, SSIM=0.829
VSNR=21.72 dB



(c) PSNR=30.24 dB, SSIM=0.832
VSNR=21.94 dB



(d) PSNR=30.56 dB, SSIM=0.837
VSNR=22.58 dB

FIG. 2 – Images “lena” reconstruites à 0.125 bpp en utilisant (b) NSLS(2,2) (c) NSLS(2,2)-OPT-L2 (d) NSLS(2,2)-OPT-WL1.

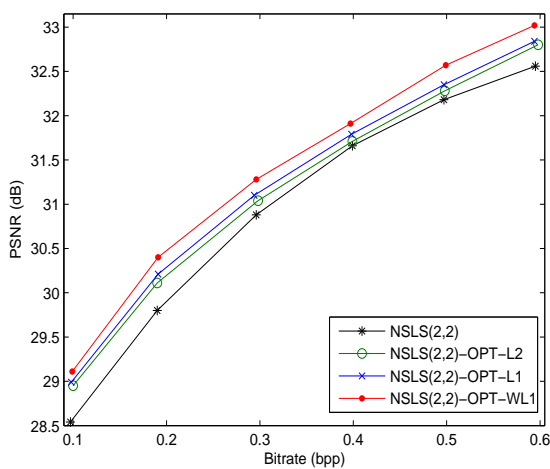


FIG. 3 – PSNR (en dB) en fonction du débit (en bpp) de l’image “einst” en utilisant le codeur entropique EBCOT.

5 Conclusion et perspectives

Dans cet article, nous avons proposé une nouvelle approche d’optimisation des différents filtres mis en jeu dans une structure de lifting 2D non séparable de type P-U. Cette approche a permis de réaliser un gain significatif en terme de qualité de reconstruction par rapport à la méthode usuelle reposant sur la minimisation d’un critère quadratique. Comme nous l’avons mentionné, la structure de lifting considérée est équivalente à toute structure de lifting classique séparable avec un seul étage de lifting. Dans le futur, nous envisageons l’extension de cette approche d’optimisation à d’autres structures de lifting ayant plus d’un étage comme les structures P-U-P [11] et P-U-P-U.

Références

- [1] W. Sweldens. The lifting scheme : a new philosophy in biorthogonal wavelet construction. In *Wavelet Applications in Signal and Image Processing III, SPIE*, pages 68–79, San-Diego, CA, USA, September 1995.
- [2] A. Gouze, M. Antonini, M. Barlaud, and B. Macq. Design of signal-adapted multidimensional lifting schemes for lossy coding. *IEEE Transactions on Image Processing*, 13(12) :1589–1603, December 2004.
- [3] V. Chappelier and C. Guillemot. Oriented wavelet transform for image compression and denoising. *IEEE Transactions on Image Processing*, 15(10) :2892–2903, October 2006.
- [4] A. Benazza-Benyahia, J.-C. Pesquet, J. Hattay, and H. Mas-moudi. Block-based adaptive vector lifting schemes for multi-channel image coding. *EURASIP International Journal of Image and Video Processing*, 2007(1) :10 pages, January 2007.
- [5] G. Jeannic, V. Ricordel, and D. Barba. The edge driven oriented wavelet transform : an anisotropic multidirectional representation with oriented lifting scheme. In *Visual Communications and Image Processing*, volume 2, San Jose, USA, 2007.
- [6] G. Piella, B. Pesquet-Popescu, H. Heijmans, and G. Pau. Combining seminorms in adaptive lifting schemes and applications to image analysis and compression. *Journal of Mathematical Imaging and Vision*, 25(2), July 2006.
- [7] M. Kaaniche, A. Benazza-Benyahia, B. Pesquet-Popescu, and J.-C. Pesquet. Non separable lifting scheme with adaptive update step for still and stereo image coding. *Elsevier Signal Processing : Special issue on Advances in Multirate Filter Bank Structures and Multiscale Representations*, January 2011.
- [8] P. L. Combettes and J.-C. Pesquet. Proximal splitting methods in signal processing. In H. H. Bauschke, R. Burachik, P. L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz, editors, *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*. Springer-Verlag, New York, 2010.
- [9] B. Usevitch. Optimal bit allocation for biorthogonal wavelet coding. In *Data Compression Conference*, pages 387–395, Snowbird, USA, March 1996.
- [10] D. M. Chandler and S. S. Hemami. VSNR : A wavelet-based visual signal-to-noise ratio for natural images. *IEEE Transactions on Image Processing*, 16(9) :2284–2298, September 2007.
- [11] M. Kaaniche, A. Benazza-Benyahia, B. Pesquet-Popescu, and J.-C. Pesquet. Vector lifting schemes for stereo image coding. *IEEE Transactions on Image Processing*, 18(11) :2463–2475, November 2009.